

# What Gives an Insulin Hexamer Its Unique Shape and Stability? Role of Ten Confined Water Molecules

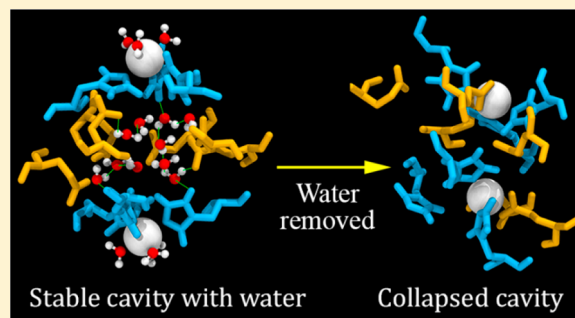
Saumyak Mukherjee,<sup>†</sup> Sayantan Mondal,<sup>†</sup> Ashish Anilrao Deshmukh,<sup>‡</sup> Balasubramanian Gopal,<sup>‡</sup> and Biman Bagchi<sup>\*,†,‡</sup>

<sup>†</sup>Solid State and Structural Chemistry Unit, Indian Institute of Science, Bangalore, India

<sup>‡</sup>Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India

## Supporting Information

**ABSTRACT:** Self-assembly of proteins often gives rise to interesting quasi-stable structures that serve important biological purposes. Insulin hexamer is such an assembly. While monomer is the biologically active form of insulin, hexamer serves as the storehouse of the hormone. The hexamer also prevents the formation of higher order aggregates. While several studies explored the role of bivalent metal ions like  $\text{Zn}^{2+}$ ,  $\text{Ca}^{2+}$ , etc., in the stabilization of the hexameric form, the role of water molecules has been ignored. We combine molecular dynamics simulations, quantum calculations, and X-ray analyses to discover that a team of approximately 10 water molecules confined inside a barrel-shaped nanocavity at the center of insulin hexamer is one of the major causes that account for the unusual stability of the biomolecular assembly. These cavity water molecules exhibit interesting dynamical features like intermittent escape and reentrance. We find that these water molecules are dynamically slower than the bulk and weave an intricate hydrogen bond network among themselves and with neighboring protein residues to generate a robust backbone at the center of the hexamer that holds the association strongly from inside and maintains the barrel shape.



## INTRODUCTION

In Nature, we find exotic structures formed by self-assembly of small organic molecules and polymers.<sup>1,2</sup> Often such self-assembled structures acquire alluring shapes and stability. In most cases, these structures are held together by hydrogen bonds, hydrophobic interactions, and coordination of oxygen and nitrogen atoms with metal ions. Study of self-assembly has blossomed into a major area of research and remains at the limelight for its immense implications in biological processes.<sup>3,4</sup> For example, protein aggregation is the cause of major ailments like Alzheimer's disease, Parkinson's disease, etc.<sup>5,6</sup>

Proteins are often found to associate into robust structures that serve important biological purposes. Insulin hexamer is such a self-assembled structure. Out of the several oligomeric forms of insulin, it is the most stable. However, monomer serves as the biologically active form.<sup>7</sup> The structure of monomeric insulin itself is interesting. It consists of two chains, A and B, with distinctly different chemical natures and solvation properties<sup>8</sup> (see Figure S1 for details). Association of two monomers leads to the formation of an insulin dimer. The major driving force of this process is provided by an extended hydrophobic patch at the junction of the two monomers (antiparallel  $\beta$ -sheets in chain-B) because of the presence of hydrophobic residues, namely, phenyl alanine (Phe-24 and Phe-25). The dimeric association is further strengthened by four hydrogen bonds (HBs) among these residues from the two monomers (Figure S2). Karplus and co-workers have shown

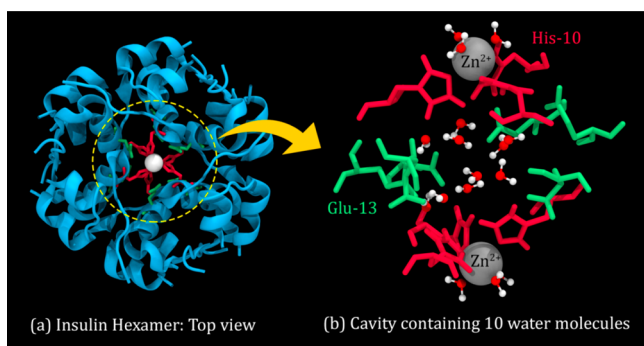
that the free energy of insulin dimerization is  $-11.9 \text{ kcal mol}^{-1}$ ,<sup>9</sup> with the experimentally determined value being  $-7.2 \text{ kcal mol}^{-1}$ .

Self-association of these dimers takes place in the presence of bivalent metal ions like  $\text{Zn}^{2+}$ ,  $\text{Ca}^{2+}$ , etc., giving rise to the highly robust and symmetric insulin hexamer structures.<sup>7,11</sup> In the human body, insulin is formed and stored in the  $\text{Zn}^{2+}$  rich vesicles in the  $\beta$ -cells of the pancreas. Six histidine residues from chain-B (His-10) of each monomer get coordinated to two  $\text{Zn}^{2+}$  ions, giving insulin hexamer its unique torus shape.<sup>12</sup> A  $C_3$  axis of rotation passes through the two  $\text{Zn}^{2+}$  ions (Figure 1a). Isothermal titration calorimetric (ITC) study by Carpenter and Wilcox shows that formation of insulin hexamer accounts for a total free energy change of  $-26 \text{ kcal mol}^{-1}$ .<sup>13</sup> Note that the stable hexamer not only acts as a storage unit but also prevents aggregation which is a major biomedical problem.<sup>14</sup>

The hexamer accommodates a barrel-shaped cavity at its center, walled on the inside by six glutamate residues and three histidine residues coordinated to one  $\text{Zn}^{2+}$  ion at the two ends (Figure 1b). From the available crystal structures and our MD simulation trajectory, we find that the average distance between the two  $\text{Zn}^{2+}$  ions is 1.4 nm, while the cross section of the barrel spans 1.1 nm. The fluctuation in the distance between the two  $\text{Zn}^{2+}$  ions is only 8%, which shows that the cavity has a

Received: January 15, 2018

Published: January 17, 2018



**Figure 1.** (a) Insulin hexamer viewed through the  $C_3$  axis of rotation (top view). (b) Insulin hexamer cavity (side view) walled by 6 glutamates (Glu-13), 6 histidines (His-10), and 2  $Zn^{2+}$  ions. It contains 10 water molecules, while 6 more are coordinated to the  $Zn^{2+}$  ions.

highly stable constitution. We find that on average the insulin hexamer cavity encapsulates 10 water molecules. This is aided by the adequate dimensions of the cavity and a polar environment provided by the hydrophilic residues like Glu-13 and His-10.

Although the roles of bivalent metal ions in stabilizing the insulin hexamer structure have been widely studied,<sup>11</sup> the role of water molecules has not been explored. Earlier experimental and simulation studies have established that water molecules are crucial to the proper functioning and structure maintenance of biomolecules.<sup>15,16</sup> Water in the hydration layer and those confined in pockets or cavities of proteins have been widely investigated.<sup>15,17–19</sup> Such confined waters are often found to play important biophysical roles. For example, proton transfer by bacteriorhodopsin is known to be mediated by water molecules trapped inside the biomolecular assembly.<sup>20,21</sup> Polar cavities inside proteins are known to house dynamically ordered water molecules.<sup>22–24</sup> Often, by the virtue of hydrogen bonding, the confined water molecules provide extra stability to the biomolecular assemblies. Hence, the natural question that comes to one's mind is the following. *What role do the water molecules inside the insulin hexamer cavity play in maintaining the structural integrity of the biomolecular assembly?*

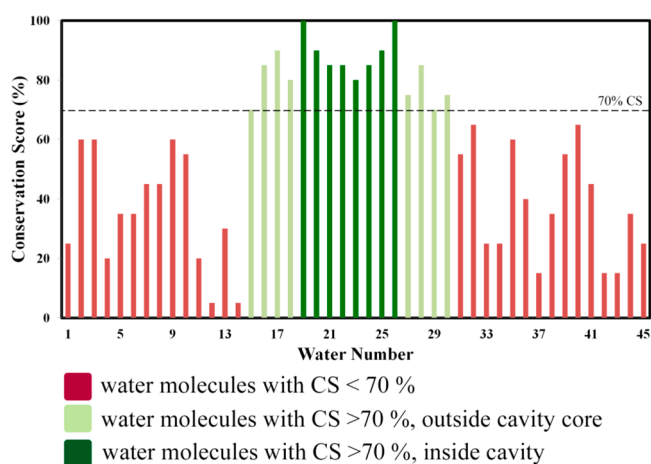
In this Article, we attempt to address this question by implementation of MD simulations, X-ray analyses, and quantum chemical calculations. We start with the analyses of conservation of cavity water in different crystal structures of insulin hexamer and the  $B$ -factors of the same from X-ray crystallography. This is followed by our observations and interpretations from MD simulation. Thereafter, we study the hydrogen bond dynamics of the cavity water molecules. Finally, to further validate our results, we perform a simulation of insulin hexamer after removing the cavity water and study the subsequent change in the structure and symmetry of the cavity.

## RESULTS AND DISCUSSION

**Conserved Water Analysis.** In most of the available PDB (Protein Data Bank) structures of insulin hexamer, water molecules are present in the cavity. This proves that water has a propensity to crystallize inside the cavity along with the hexamer itself. We superimpose 20 such structures (solved by high resolution X-ray crystallography) (Table S1) in COOT<sup>25</sup> to calculate their conservation scores (CS) (eq 1).

$$CS (\%) = \frac{n}{20} \times 100 \quad (1)$$

Here,  $n$  is the number of water molecules present at a particular position corresponding to all of the 20 structures. Hence, a higher conservation score signifies that the water molecule has a higher tendency to remain at that position in different crystal structures. The subsequent results are plotted in Figure 2.



**Figure 2.** Conservation scores of water molecules in an insulin hexamer. Dark green lines represent water molecules inside the core of the cavity. Light green lines are for the molecules which are hydrogen bonded to protein residues but do not reside at the core of the cavity. Red lines indicate water outside of the cavity.

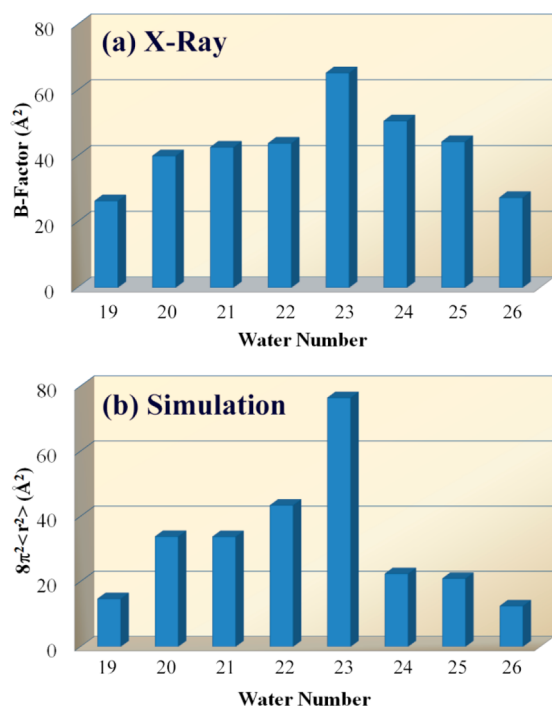
Water molecules that reside at the core of the cavity (19–26) have conservation scores  $>70\%$  (dark green). They are hydrogen bonded to surrounding amino acid side chains. The molecules that are coordinated to the two  $Zn^{2+}$  ions are totally conserved (19, 26) ( $CS = 100\%$ ); that is, they are present in all of the crystal structures. There are certain other conserved water molecules (light green, 15–18 and 27–30) that reside at the junction of two dimers in the hexameric association and might play a significant role in holding them together. The molecules represented by red lines are outside the cavity and have lower CS values (1–14, 31–45). This analysis thus demonstrates that water molecules inside the cavity are localized and therefore are crucial in effecting the stability of the cavity.

**$B$ -Factors of Cavity Water.** The spatial rigidity of these localized water molecules is further established by calculating  $B$ -factors (eq 2),<sup>26</sup> which is widely used to quantify the rigidity of an atom or part of a system in crystallography.

$$B = 8\pi^2 \langle r^2 \rangle \quad (2)$$

Here,  $r^2$  is the mean square displacement (MSD) of the moiety of interest. In Figure 3, we show the  $B$ -factors of cavity waters calculated from X-ray crystallography and its analogue from MD simulation, respectively, both being consistent with each other.

Molecules 19 and 26, the ones which are coordinated to  $Zn^{2+}$  ions, have the least  $B$ -factor values, while the central water (23) has the maximum value. This is because, being coordinated to  $Zn^{2+}$  ions, molecules 19 and 26 are strongly localized, while molecule 23 has a greater freedom to move around. Other water molecules, which are hydrogen bonded to nearby protein residues, also have lower values of  $B$ -factor compared to bulk water (which is often of the order of  $100 \text{ \AA}^2$ ). This again depicts the spatial restriction of the cavity waters that provide a firm backbone stabilizing the hexameric association.



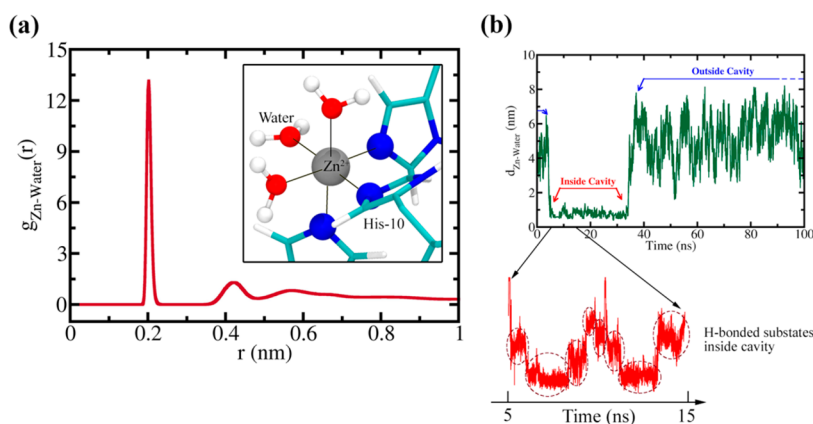
**Figure 3.** B-Factor analysis of cavity waters from (a) X-ray crystallography and (b) analogous analysis in computer simulation.

**Observations from Computer Simulation.** While X-ray furnishes the coordinates for a single conformation, MD simulation gives results that are ensemble averaged. The eight cavity water molecules presented in the X-ray analysis correspond to the asymmetric unit of insulin hexamer. Hence, on application of  $C_3$  symmetry to visualize the whole hexamer (biological unit), the number of water molecules that do not lie on the rotation axis (five) multiplies and results in a total of 18 cavity water molecules. However, molecular simulation gives a more realistic picture of the actual number of water molecules that the cavity can accommodate. From our 100 ns MD trajectory, we find that the insulin hexamer cavity can house 10 water molecules at its core, on average. It also provides evidence for the existence of  $Zn^{2+}$ -coordinated water

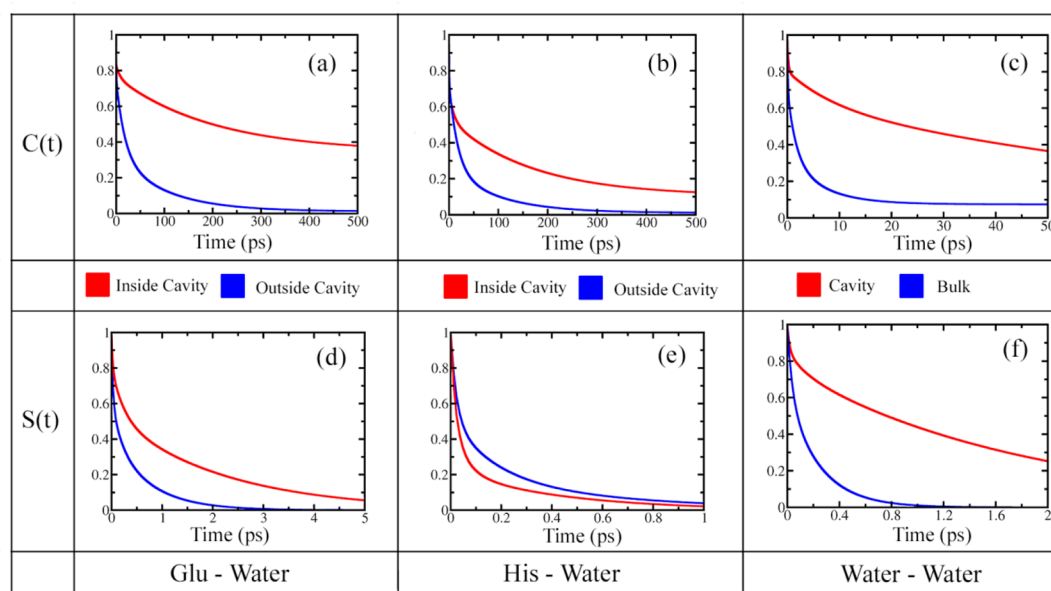
molecules. We find that three water molecules remain coordinated to each of the two  $Zn^{2+}$  ions (Figure 1b). The radial distribution function (RDF) between  $Zn^{2+}$  and the oxygen atom of water molecules shows a sharp peak at 0.2 nm corresponding to these coordinated water molecules (Figure 4a). To further establish this fact, we monitor the distance of these water molecules from the  $Zn^{2+}$  ions as a function of time. This trajectory shows minimal fluctuations ( $\sim 0.2$  nm) that substantiate the RDF peak.

These water molecules, along with three His-10 residues per  $Zn^{2+}$  ion, create an octahedral coordination shell around the ion (Figure 4a inset).<sup>27–29</sup> However, reports of tetrahedral arrangements (with three His-10 and one water) are also available in the literature.<sup>30–32</sup> In our simulations, the octahedral geometry is preserved throughout the 100 ns trajectory, indicating a very stable coordination geometry. Thus, we ask the question, how stable is this complex? To quantify the relative stability of this complex as opposed to hydrated  $Zn^{2+}$  ions, we perform quantum DFT calculations in the Gaussian 09<sup>33</sup> package. The B3LYP<sup>34</sup> functional and 6-311G+(d,p) basis set are used for energy calculation. We probe the stability of the hexamer by considering the complex formation among  $Zn^{2+}$  ions, imidazole nitrogens (from His-10 residues), and water molecules. We calculate the energies of individual histidine, water, and  $Zn^{2+}$  along with the energies of other possible complexes. The geometry of the complex here is fixed according to the average equilibrated structure of insulin hexamer in water as obtained from simulation. Basis set superposition error (BSSE) is calculated using counterpoise<sup>35</sup> for every structure. We find that a  $[Zn(H_2O)_3(Im)_3]^{2+}$  (“Im” = imidazole, present in histidine side chain) complex is 12.6 kcal mol<sup>−1</sup> more stable than a  $[Zn(H_2O)_6]^{2+}$  complex (which is the octahedral coordination environment of a  $Zn^{2+}$  ion in aqueous solution). Such a stable coordination environment ensures a robust structure at the two ends of the barrel shaped hexamer cavity.

Although the coordinated water molecules are not exchanged during the entire span of our simulation trajectory (100 ns), certain molecules exhibit intermittent escape and reentrance into the cavity. A similar type of flow of water molecules was



**Figure 4.** (a) Pair correlation function between  $Zn^{2+}$  ions and water oxygen [ $g_{Zn-water}(r)$ ]. The sharp peak at 0.20 nm corresponds to the three coordinated water molecules that create an octahedral atmosphere around each  $Zn^{2+}$  ion along with three His-10 residues (inset). (b) Distance of a tagged water molecule from a  $Zn^{2+}$  ion ( $d_{Zn-water}$ ) for a 100 ns trajectory. From 5 to 35 ns, the water resides inside the cavity, denoted by a steady decreased distance from the  $Zn^{2+}$  ion. Inside the cavity, the water molecule attains certain stable H-bonded substates (circled by dashed lines) (represented for a duration of 10 ns here).



**Figure 5.** HB time correlation functions of water molecules among themselves and nearby protein residues. (a)  $C(t)$  for glutamate–water HB; (b)  $C(t)$  for histidine–water HB; (c)  $C(t)$  for water–water HB; (d)  $S(t)$  for glutamate–water HB; (e)  $S(t)$  for histidine–water HB; (f)  $S(t)$  for water–water HB. Red curves are for the dynamics inside the cavity, and blue curves represent the same outside or in the bulk.

reported earlier by Hummer et al. in the case of carbon nanotube systems.<sup>36,37</sup>

Figure 4b shows the distance trajectory of such a water molecule from a  $\text{Zn}^{2+}$  ion ( $d_{\text{Zn}^{2+}\text{-water}}$ ). This particular water enters the cavity at 5 ns and leaves it at 35 ns. From Figure 3b, it becomes clear that the molecule gets trapped inside the cavity with a long residence time ( $\sim 30$  ns). Confined waters show slower dynamics than bulk water.<sup>38</sup> The time taken by cavity waters to suffer a displacement equal to one molecular diameter (3.166 Å for SPC/E) is often above 50 ps (sometimes as high as 150 ps). However, bulk water takes only 5 ps on average. Relaxation of rotational and total dipole moment time correlation functions is also slower inside the cavity (sections S5 and S6 in the Supporting Information). From Figure 4b, we see that the fluctuation of the  $\text{Zn}^{2+}$ –water distance in bulk is about 6 nm, while inside the cavity it is only  $\sim 0.5$  nm. Inside the cavity, the water molecule resides in certain substates that can be explained in terms of long-lived hydrogen bonds (HBs). This hydrogen bonding is the most crucial tool that allows cavity waters to reinforce the insulin hexamer assembly. Similar kinds of trajectories are shown in Figure S3 for other water molecules.

However, we note that the characteristics of the cavity are significantly affected by the presence of other ions. When  $\text{Ca}^{2+}$  ions are present in the environment during hexamer formation, one  $\text{Ca}^{2+}$  is found to get coordinated to the Glu-13 residues at the center of the cavity.<sup>11</sup> Occupation of the cavity by a  $\text{Ca}^{2+}$  ion results in reduced available space in the cavity, affecting the number of water molecules and thus the hydrogen bond network. However, in our study, we focus on insulin hexamer with only two  $\text{Zn}^{2+}$  ions and the subsequent cavity characteristics.

**Hydrogen Bond Analysis.** While HBs are pivotal to any biomolecular assembly, the HB network spanned by the cavity waters draws special interest. In order to study the dynamics of this HB network, we define two time correlation functions (TCFs), namely, the intermittent HB TCF ( $C(t)$ ) and the continuous HB TCF ( $S(t)$ ) given by the following eq 3.<sup>39,40</sup>

$$C(t) = \frac{\langle h(0)h(t) \rangle}{\langle h \rangle} \quad S(t) = \frac{\langle h(0)H(t) \rangle}{\langle h \rangle} \quad (3)$$

Here  $h(t)$  is a population parameter which attains a value of “1” when a particular H-bond exists at time  $t$  and “0” otherwise.  $H(t)$  is a similar parameter which has a value of “1” as long as a H-bond exists and becomes “0” for the rest of the trajectory when it breaks for the first time. Hence,  $C(t)$  provides us with the overall structural information on the HB network, whereas  $S(t)$  estimates its lifetime. Details of the theoretical definition of the H-bond are discussed in section S4 of the Supporting Information.

In Figure 5, red curves represent HB dynamics for the cavity and blue curves represent the same for regions outside of the cavity.  $C(t)$  for water–Glu and water–His H-bonds are shown in parts a and b of Figure 5, respectively. In both cases, relaxation is much slower inside the cavity. It is approximately 3.5 times slower in the case of water–Glu HB and 2.5 times slower in the case of water–His HB (Table 1). This slower relaxation of  $C(t)$  inside the cavity results from localized water molecules and signifies a well-formed HB network inside the cavity.

However, the scenario is a little different in the case of  $S(t)$ . Relaxation of this continuous HB TCF is slower for water–Glu H-bond but faster for water–His H-bond inside the cavity in contrast to those outside. This difference can be attributed to the spatial constraints faced by the coordination of His-10 with  $\text{Zn}^{2+}$  ions. Owing to this coordination, the side chains of His-10 are not free to move along with the movement of water molecules. This results in recurrent breaking and formation of the corresponding H-bonds, resulting in lower lifetimes. On the other hand, Glu-13 side chains in the cavity are not constrained, providing the Glu–water H-bonds a greater lifetime. This difference reflects the cooperativity between amino acid side chains and water molecules in maintaining the long lifetime of these H-bonds.

Other than the hydrogen bonding among water and amino acid side chains, the 10 cavity waters themselves contribute

**Table 1.** Average Relaxation Times for HB Dynamics of Water and Neighboring Residues inside and outside the Cavity

HB TCF	HB of water with	average relaxation time (ps)
$C(t)$	Glu-13 ( <i>cavity</i> )	141
	Glu-4,17,21 ( <i>outside</i> )	40
	His-10 ( <i>cavity</i> )	82
	His-5 ( <i>outside</i> )	33
	water ( <i>cavity</i> )	110.35
	water ( <i>bulk</i> )	2.52
$S(t)$	Glu-13 ( <i>cavity</i> )	1.240
	Glu-4,17,21 ( <i>outside</i> )	0.350
	His-10 ( <i>cavity</i> )	0.120
	His-5 ( <i>outside</i> )	0.184
	water ( <i>cavity</i> )	1.39
	water ( <i>bulk</i> )	0.168

significantly to the HB network. On average, there are 15 hydrogen bonds among these 10 molecules. The dynamics of these hydrogen bonds are much slower than that for an equal number of water molecules in bulk water (Figure 5c, Figure 5f, and Table 1). Consequently, this maintains the continuity of the HB network inside the insulin hexamer cavity, thereby providing a robust backbone to the biomolecular association.

**Fate of the Cavity in the Absence of Water.** The analyses presented thus far suggest that cavity water is responsible for the stability of insulin hexamer. In order to further substantiate this observation, we monitor the structure of the cavity in the absence of water in computer simulation (note that the initial structure of this simulation has no water molecules in the cavity; however, water is present outside the cavity). Analysis of the trajectory reveals that, when water is removed, the cavity breaks down within a few picoseconds. The  $Zn^{2+}$  ions come closer to an average distance of 0.8 nm (that was  $\sim 1.4$  nm in the presence of water). In the absence of water molecules, coordination sites on  $Zn^{2+}$  become vacant and Glu-13 residues with their bare negatively charged side chains coordinate with the  $Zn^{2+}$  ions. This helps in reducing the repulsive interactions among the negatively charged carboxylate moieties. We observe that two Glu-13 residues get coordinated to each  $Zn^{2+}$  ion, thereby altering the  $Zn^{2+}$ –His-10 coordination (Figure 6). These sequential steps lead to the collapse of the cavity disrupting the symmetric arrangement of the insulin hexamer into an orderless aggregate. The role of the conserved water molecules in initiating this dissociation suggests a dominant role of hydration forces in the hexameric insulin assembly.

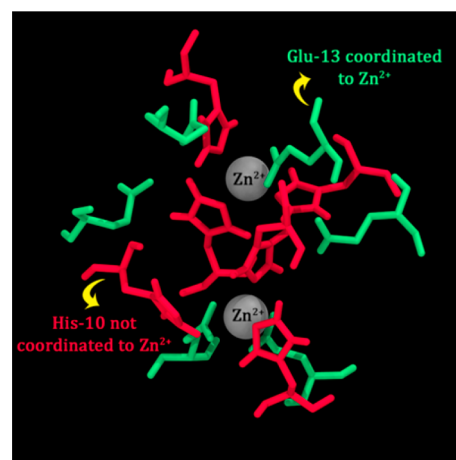
## CONCLUSION

Let us summarize the main results of this paper.

(i) We have combined computer simulation, X-ray crystallography, and quantum calculations to establish, for the first time, the role of cavity water molecules in stabilizing the insulin hexamer.

(ii) We find that these localized water molecules efficiently screen the electrostatic interactions among the charged moieties inside the cavity.

(iii) In addition, they weave a strong HB network at the middle of the hexameric association of insulin that holds it from inside. This robust HB network endows the hexamer with an extraordinary stability, as manifested in the low fluctuation of



**Figure 6.** Collapse of the insulin hexamer cavity in the absence of water. When water molecules are absent, the charges on Glu-13 side chains are no longer screened and they get coordinated to  $Zn^{2+}$  ions. This leads to the demolition of the structural integrity of the hexamer, rendering it into an aggregation devoid of any symmetry.

$Zn$ – $Zn$  distance, which, on an average, is found to be only 8% of the average separation.

(iv) It is well-known that conformations and structural aspects of biomolecules are pivotal to their biological functions.<sup>41</sup> Our study demonstrates that, in the absence of these confined water molecules, the cavity collapses, which results in structural changes in the hexamer and consequently affects its biological functions.

Although our study focuses on a particular protein, such a role of water molecules can be universal and may be important in similar kinds of self-assemblies. These results have implications in a broader perspective of the microscopic role played by water molecules in rendering an unusual stability to biomolecular assemblies by means of an extraordinarily intricate hydrogen bond network that works as a framework, defining the structural integrity of the system.

## EXPERIMENTAL SECTION

Insulin was obtained from Sigma-Aldrich, and initial crystallization trials were performed using commercially available crystallization screen from Hampton research. This hormone was crystallized (0.1 M sodium acetate trihydrate, pH 4.6, 2.0 M sodium chloride) at  $7 \text{ mg mL}^{-1}$  insulin containing a trace amount of zinc chloride. Crystals were obtained and soaked with 10% ethylene glycol in mother liquor prior to data collection. Data was collected at home source and processed by iMOSFLM<sup>42</sup> and scaled using SCALA.<sup>43</sup> The phase information was obtained by the molecular replacement method using the insulin model (PDB code: 3W7Y). The model was refined using REFMAC5,<sup>44</sup> and the fit of the model to the electron density was evaluated using COOT.<sup>37</sup> Data collection and refinement statistics for insulin hexamer and PDB validation statistics are presented in section S8 of the Supporting Information.

## COMPUTATIONAL SECTION

**Molecular Dynamics Simulation.** We performed atomistic molecular dynamics (MD) simulations using the GROMACS-4.5.6 package<sup>45</sup> which is a widely accepted and highly efficient MD engine. The initial configuration of the system was taken from the crystal structure available in the

Protein Data Bank (PDB code: 3W7Y). The asymmetric unit in 3W7Y was processed to get the biological unit that is hexamer, using UCSF Chimera-1.11.2.<sup>46</sup> For simulation, we used the GROMOS96 53a6<sup>47</sup> force field for protein and the extended simple point charge model (SPC/E)<sup>48</sup> for water. Periodic boundary conditions were implemented using a cubic box of 10 nm dimensions with 31 620 water molecules in the system.

The total system was energy minimized by a succession of steepest descent and conjugate gradient algorithms.<sup>49</sup> Thereafter, the solvent (water) was equilibrated under NPT conditions ( $T = 3300$  K and  $P = 1$  bar), restraining the positions of protein atoms for 5 ns followed by a similar equilibration under NVT conditions ( $T = 300$  K). Then, the system was subjected to a further 10 ns NVT equilibration ( $T = 300$  K) without any position restraints. The final production run was carried out in an NVT environment at a temperature of 300 K for 105 ns. The last 100 ns were taken for analyses. Data was dumped at a frequency of 1.0 ps for analyzing static properties. Another MD run of 1 ns was performed, and data was saved at 2 fs resolution for analyses of dynamic properties. The equations of motion were integrated using the leap-frog algorithm with a time step of 1 fs.

In order to maintain a fixed average temperature and pressure, we use a Nosé–Hoover thermostat<sup>50</sup> ( $\tau_t = 0.1$  ps and two coupling groups, namely, protein and nonprotein) and a Parrinello–Rahman barostat<sup>51</sup> ( $\tau_p = 2.0$  ps), respectively. A cutoff radius of 10 Å was set for neighbor searching and calculation of nonbonded interactions, and all bonds were constrained using the LINCS algorithm.<sup>52</sup> For calculation of electrostatic interactions, the particle mesh Ewald method<sup>53</sup> was used with a FFT grid spacing of 1.6 Å.

## ■ ASSOCIATED CONTENT

### ● Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jpcc.8b00453.

Structures of insulin monomer and dimer, details of PDB structures used in conserved water analysis,  $\text{Zn}^{2+}$ –water distance trajectories, theoretical definition of hydrogen bond, rotational dynamics, total dipole moment correlation function, density of states, data collection, and refinement statistics of insulin hexamer (PDF)

Video showing collapse of the cavity in absence of water (MPG)

## ■ AUTHOR INFORMATION

### Corresponding Author

\*E-mail: profbiman@gmail.com.

### ORCID

Biman Bagchi: 0000-0002-7146-5994

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

We thank Prof. P. Balaram and Prof. P. Wolynes for insightful discussions. We thank the Department of Science and Technology (DST, India) for partial support of this work. B.B. thanks Sir J. C. Bose fellowship for partial support. S.Mo. and A.A.D. thank UGC, India, for providing a research fellowship. S.Mu. thanks DST, India, for providing an INSPIRE fellowship.

## ■ REFERENCES

- (1) Lehn, J.-M. *Supramolecular chemistry*; VCH: Weinheim, Germany, 1995; Vol. 1.
- (2) Palmer, B. J.; Liu, J. Simulations of micelle self-assembly in surfactant solutions. *Langmuir* **1996**, *12* (3), 746–753.
- (3) Ferreira, F. M.; Oliveira, L. C.; Germino, G. G.; Onuchic, J. N.; Onuchic, L. F. Macromolecular assembly of polycystin-2 intracytosolic C-terminal domain. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (24), 9833–9838.
- (4) Dima, R. I.; Thirumalai, D. Exploring protein aggregation and self-propagation using lattice models: Phase diagram and kinetics. *Protein Sci.* **2002**, *11* (5), 1036–1049.
- (5) Thirumalai, D.; Reddy, G.; Straub, J. E. Role of Water in Protein Aggregation and Amyloid Polymorphism. *Acc. Chem. Res.* **2012**, *45* (1), 83–92.
- (6) Chiti, F.; Dobson, C. M. Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem.* **2006**, *75*, 333–366.
- (7) Blundell, T.; Dodson, G.; Hodgkin, D.; Mercola, D. Insulin: The Structure in the Crystal and its Reflection in Chemistry and Biology by. *Adv. Protein Chem.* **1972**, *26*, 279–402.
- (8) Bagchi, K.; Roy, S. Sensitivity of water dynamics to biologically significant surfaces of monomeric insulin: Role of topology and electrostatic interactions. *J. Phys. Chem. B* **2014**, *118* (14), 3805–3813.
- (9) Zoete, V.; Meuwly, M.; Karplus, M. Study of the insulin dimerization: Binding free energy calculations and per-residue free energy decomposition. *Proteins: Struct., Funct., Genet.* **2005**, *61* (1), 79–93.
- (10) Strazza, S.; Hunter, R.; Walker, E.; Darnall, D. W. The thermodynamics of bovine and porcine insulin and proinsulin association determined by concentration difference spectroscopy. *Arch. Biochem. Biophys.* **1985**, *238* (1), 30–42.
- (11) Dunn, M. F. Zinc–ligand interactions modulate assembly and stability of the insulin hexamer—a review. *BioMetals* **2005**, *18* (4), 295–303.
- (12) Hill, C. P.; Dauter, Z.; Dodson, E. J.; Dodson, G. G.; Dunn, M. F. X-ray structure of an unusual calcium site and the roles of zinc and calcium in the assembly, stability, and storage of the insulin hexamer. *Biochemistry* **1991**, *30* (4), 917–924.
- (13) Carpenter, M. C.; Wilcox, D. E. Thermodynamics of Formation of the Insulin Hexamer: Metal-Stabilized Proton-Coupled Assembly of Quaternary Structure. *Biochemistry* **2014**, *53* (8), 1296–1301.
- (14) Lisi, G. P.; Png, C. Y. M.; Wilcox, D. E. Thermodynamic contributions to the stability of the insulin hexamer. *Biochemistry* **2014**, *53* (22), 3576–3584.
- (15) Mondal, S.; Mukherjee, S.; Bagchi, B. Protein Hydration Dynamics: Much Ado about Nothing? *J. Phys. Chem. Lett.* **2017**, *8* (19), 4878–4882.
- (16) Levy, Y.; Onuchic, J. N. Water mediation in protein folding and molecular recognition. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35* (1), 389–415.
- (17) Bagchi, B. *Water in Biological and Chemical Processes: From Structure and Dynamics to Function*; Cambridge University Press: Cambridge, UK, 2013.
- (18) Vajda, S.; Jimenez, R.; Rosenthal, S. J.; Fidler, V.; Fleming, G. R.; Castner, E. W. Femtosecond to nanosecond solvation dynamics in pure water and inside the [gamma]-cyclodextrin cavity. *J. Chem. Soc., Faraday Trans.* **1995**, *91* (5), 867–873.
- (19) Biswas, R.; Furtado, J.; Bagchi, B. Layerwise decomposition of water dynamics in reverse micelles: A simulation study of two-dimensional infrared spectrum. *J. Chem. Phys.* **2013**, *139* (14), 144906.
- (20) Kandori, H. Role of internal water molecules in bacteriorhodopsin. *Biochim. Biophys. Acta, Bioenerg.* **2000**, *1460* (1), 177–191.
- (21) Grudinin, S.; Büldt, G.; Gordeliy, V.; Baumgaertner, A. Water Molecules and Hydrogen-Bonded Networks in Bacteriorhodopsin—Molecular Dynamics Simulations of the Ground State and the M-Intermediate. *Biophys. J.* **2005**, *88* (5), 3252–3261.
- (22) Quillin, M. L.; Wingfield, P. T.; Matthews, B. W. Determination of solvent content in cavities in IL-1 $\beta$  using experimentally phased

electron density. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103* (52), 19749–19753.

(23) Yin, H.; Feng, G.; Clore, G. M.; Hummer, G.; Rasaiah, J. C. Water in the Polar and Nonpolar Cavities of the Protein Interleukin-1 $\beta$ . *J. Phys. Chem. B* **2010**, *114* (49), 16290–16297.

(24) Liu, L.; Quillin, M. L.; Matthews, B. W. Use of experimental crystallographic phases to examine the hydration of polar and nonpolar cavities in T4 lysozyme. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (38), 14406–14411.

(25) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. Features and development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66* (4), 486–501.

(26) Frauenfelder, H. The Debye-Waller factor: From villain to hero in protein crystallography. *Int. J. Quantum Chem.* **1989**, *35* (6), 711–715.

(27) Baker, E. N.; Blundell, T. L.; Cutfield, J. F.; Cutfield, S. M.; Dodson, E. J.; Dodson, G. G.; Hodgkin, D. M. C.; Hubbard, R. E.; Isaacs, N. W.; Reynolds, C. D. The Structure of 2Zn Pig Insulin Crystals at 1.5 Å Resolution. *Philos. Trans. R. Soc., B* **1988**, *319* (1195), 369–456.

(28) Smith, G.; Duax, W.; Dodson, E.; Dodson, G.; De Graaf, R.; Reynolds, C. The structure of Des-Phe B1 bovine insulin. *Acta Crystallogr., Sect. B: Struct. Crystallogr. Cryst. Chem.* **1982**, *38* (12), 3028–3032.

(29) Lisgarten, D. R.; Palmer, R. A.; Loble, C. M.; Naylor, C. E.; Chowdhry, B. Z.; Al-Kurdi, Z. I.; Badwan, A. A.; Howlin, B. J.; Gibbons, N. C.; Saldanha, J. W. Ultra-high resolution X-ray structures of two forms of human recombinant insulin at 100 K. *Chem. Cent. J.* **2017**, *11* (1), 73.

(30) Ciszak, E.; Smith, G. D. Crystallographic evidence for dual coordination around zinc in the T3R3 human insulin hexamer. *Biochemistry* **1994**, *33* (6), 1512–1517.

(31) Smith, G. D.; Ciszak, E.; Magrum, L. A.; Pangborn, W. A.; Blessing, R. H. R6 hexameric insulin complexed with m-cresol or resorcinol. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2000**, *56* (12), 1541–1548.

(32) Palmieri, L. C.; Fávero-Retto, M. P.; Lourenço, D.; Lima, L. M. T. A T3R3 hexamer of the human insulin variant B28Asp. *Biophys. Chem.* **2013**, *173–174*, 1–7.

(33) Frisch, M.; Trucks, G.; Schlegel, H.; Scuseria, G.; Robb, M.; Cheeseman, J.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. *Gaussian 09*, revision A; Gaussian, Inc.: Wallingford, CT, 2009.

(34) Stephens, P.; Devlin, F.; Chabalowski, C.; Frisch, M. J. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.* **1994**, *98* (45), 11623–11627.

(35) Boys, S. F.; Bernardi, F. The calculation of small molecular interactions by the differences of separate total energies. Some procedures with reduced errors. *Mol. Phys.* **1970**, *19* (4), 553–566.

(36) Berezhkovskii, A.; Hummer, G. Single-file transport of water molecules through a carbon nanotube. *Phys. Rev. Lett.* **2002**, *89* (6), 064503.

(37) Hummer, G.; Rasaiah, J. C.; Noworyta, J. P. Water conduction through the hydrophobic channel of a carbon nanotube. *Nature* **2001**, *414* (6860), 188–190.

(38) Bhattacharyya, K. Solvation Dynamics and Proton Transfer in Supramolecular Assemblies. *Acc. Chem. Res.* **2003**, *36* (2), 95–101.

(39) Luzar, A.; Chandler, D. Hydrogen-bond kinetics in liquid water. *Nature* **1996**, *379* (6560), 55–57.

(40) Bandyopadhyay, S.; Chakraborty, S.; Bagchi, B. Secondary structure sensitivity of hydrogen bond lifetime dynamics in the protein hydration layer. *J. Am. Chem. Soc.* **2005**, *127* (47), 16660–16667.

(41) Privalov, P. L. Stability of proteins small globular proteins. *Adv. Protein Chem.* **1979**, *33*, 167–241.

(42) Batty, T. G. G.; Kontogiannis, L.; Johnson, O.; Powell, H. R.; Leslie, A. G. W. iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67* (4), 271–281.

(43) Evans, P. Scaling and assessment of data quality. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, *62* (1), 72–82.

(44) Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. REFMAC5 for the refinement of macromolecular crystal structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67* (4), 355–367.

(45) Van Der Spoel, D.; Lindahl, E.; Hess, B.; Groenhof, G.; Mark, A. E.; Berendsen, H. J. C. GROMACS: Fast, flexible, and free. *J. Comput. Chem.* **2005**, *26* (16), 1701–1718.

(46) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25* (13), 1605–1612.

(47) Oostenbrink, C.; Villa, A.; Mark, A. E.; Van Gunsteren, W. F. A biomolecular force field based on the free enthalpy of hydration and solvation: The GROMOS force-field parameter sets 53A5 and 53A6. *J. Comput. Chem.* **2004**, *25* (13), 1656–1676.

(48) Berendsen, H.; Grigera, J.; Straatsma, T. The missing term in effective pair potentials. *J. Phys. Chem.* **1987**, *91* (24), 6269–6271.

(49) Press, W. H. *Numerical recipes 3rd edition: The art of scientific computing*; Cambridge University Press: Cambridge, UK, 2007.

(50) Hoover, W. G. Canonical dynamics: equilibrium phase-space distributions. *Phys. Rev. A: At., Mol., Opt. Phys.* **1985**, *31* (3), 1695.

(51) Parrinello, M.; Rahman, A. Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **1981**, *52* (12), 7182–7190.

(52) Hess, B.; Bekker, H.; Berendsen, H. J.; Fraaije, J. G. LINCS: a linear constraint solver for molecular simulations. *J. Comput. Chem.* **1997**, *18* (12), 1463–1472.

(53) Darden, T.; York, D.; Pedersen, L. Particle mesh Ewald: An N-log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **1993**, *98* (12), 10089–10092.